# Interaction Histories: From Experience to Action and Back Again *

Naeem Assif Mirza, Chrystopher L. Nehaniv, Kerstin Dautenhahn, René te Boekhorst
*Adaptive Systems Research Group, School of Computer Science*
*University of Hertfordshire, College Lane, Hatfield, AL10 9AB. United Kingdom*
*{N.A.Mirza,C.L.Nehaniv,K.Dautenhahn,R.teBoekhorst}@herts.ac.uk*

***Abstract*— We describe an enactive, situated model of interaction history based around a growing, informationally self-structured metric space of experience that is constructed and reconstructed as the robot engages in sensorimotor interactions with objects and people in its environment. The model shows aspects of development and learning through modification of the cognitive structure that forms the basis for action selection as a result of acting in the world. We describe robotic experiments showing prediction of the path of a ball and an interaction game "peekaboo".**

***Index Terms*— Interaction History, Information Theory, Robotic Control Architectures**

## I. Introduction

A challenge of research into situated, enactive cognition in robots is to reach beyond reactive architectures to architectures that can reflect the time-extended behaviour characteristic of humans and many animals. We are interested in how cognitive structures in natural and artificial systems can arise that capture the history of interactions and behaviours of an agent actively engaged in its environment, without resorting to symbolic representations of past events.

We introduce an architecture that has at its heart a changing dynamic structure describing the space of experience of the agent or robot. The robot chooses how to behave in the world based on what it has experienced, and this results in further experience and modification of previous experience establishing a tight coupling of experience and action.

This paper proceeds by presents our concept of an interaction history and then describes the model and architecture that we use. Finally we describe experiments conducted on a robot platform that investigate the capabilities of the model.

## II. Interaction History

We use a working definition of an *interaction history* as *the temporally extended, dynamically constructed and reconstructed, individual sensorimotor history of an agent situated and acting in its environment including the social environment.* The first key part of the definition is that the agent is situated and actively acting within its environment, that is the history is not a disembodied memory, but an active part of the interaction of the agent and its environment. This follows the idea of structural coupling and enactive cognition of Maturana and Varela [1] and the concept of situated cognition [2]. Remembering is then the effect of historical interactions on the actions of an agent in response to a particular situations [3]. This brings in the next key part of the definition, that the history is dynamically constructed and reconstructed. In other words, interactions with the environment construct the structures that are used for remembering how to act. Thus, memory consists not of static representations of the past that can be recalled with perfect clarity, but rather is the result of an accumulation of interaction with the environment manifesting as current action.

An important aspect of the interaction history is that it is constructed from the perspective of the individual, that is, it is autobiographical in nature. In terms of the accepted separation of memory types due to Endel Tulving [4], this would be episodic memory as opposed to semantic memory. That is, it is the memory of events (with a temporal aspect and, usually, a personal aspect), rather than the memory of knowledge and categories. However this apparently clear dichotomy is not applicable to a description of interaction history as, through the process of reconstruction, categories and knowledge may emerge from many overlapping experiences, while certain unique events may still stand out and give memory its episodic nature. While we do not claim that an interaction history can describe all aspects of (human) memory, we believe that exploring the features of an interaction history may give insights into the nature of memory as a whole. The final part of the definition that we would highlight is that it need not be representational but must be grounded in the sensorimotor experience of the agent.

### A. Extended Temporal Horizon

A robotic agent with an interaction history has the potential to act on an extended temporal horizon [5] resulting in behaviour that goes beyond that of a reactive agent or an affective agent. The distinction is that behaviour will be

*The work described in this paper was conducted within the EU Integrated Project RobotCub ("Robotic Open-architecture Technology for Cognition, Understanding, and Behaviours") and was funded by the European Commission through the E5 Unit (Cognition) of FP6-IST under Contract FP6-004370.

modulated by temporally extended past experience as well as by internal state (affect) and immediately by environmental stimuli (reactivity).

### *B. Development and Learning*

A further aspect of an interaction history which manifests itself as modification of behaviour based on a history of previous interactions is that it can serve to scaffold learning and development of a situated agent. The key here is how previous experience is used to affect current and future behaviour. For example, classical conditioning or a two-process reinforcement learning based on positive and negative reinforcers [6] are potential mechanisms for connecting previous experience with choice of action. Development can be seen as the increasing richness of the connections of experience with action, again mediated by a suitable mechanisms.

## III. Enactive Robot Model of Interaction History using Sensorimotor Experience

We describe a computational robotic model (Fig. 1) that illustrates how an interaction history can be integrated into the control of a robot using the concepts described in the previous section.

The basic architecture consists of processes to acquire sensory and motor values from the robot as it acts in the environment, from this a metric space of past interaction experiences is constructed. A further process continuously examines current experience in the context of the space of previous experience and selects actions to execute.

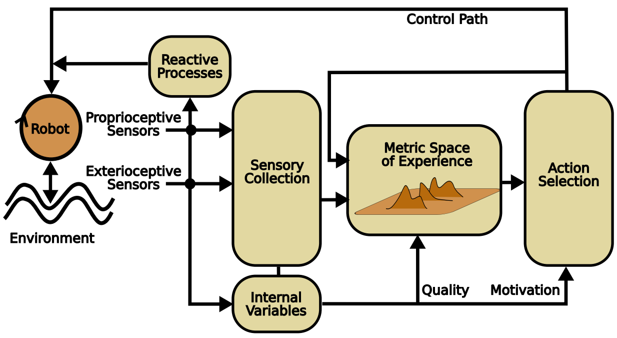


Fig. 1. Interaction history based control architecture.

### *A. Sensory and Internal Variables*

The sensory information available to the robot falls into three broad categories: proprioceptive, exteroceptive and internal. Proprioceptive variables are constructed by sampling motor position and exteroceptive variables are those from sensors such as buttons, infra-red distance, vision[1] and audition[2]. In addition to these, sensory input can also be built from internal variables that might, for instance, indicate drives and motivations, or be the result of processing of raw sensory data e.g. ball position. Sampling is done at regular intervals (between 100-120ms in the experiments presented).

### *B. Experience Space*

The experience space is constructed from overlapping experiences of a particular horizon size with relative positions in the space determined by the informational distance between them (see section *IV*). Many potential experience spaces of different horizon length can be built and co-exist [7].[3].
As the metric landscape of experience is built, each experience is further enhanced with *value attributes* of the experience. These are the instantaneous values of any sensor or internal variable, for example variables indicating "satiation", "battery-level", "contentment" and so forth. Experiences are also annotated with the actions that the robot takes at any timestep (see section *III-C*).

### *C. Action Selection, Development and Learning*

While an experience space can be built without much difficulty, the challenge is how to have experience modulate future action in a meaningful way and to be further shaped by that action. To achieve this goal, a simple mechanism is adopted whereby the robot can execute one of a number of "atomic" actions (or no action) at any timestep[4]. At any timestep the robot can choose an action based on past experience or, if an appropriate one is not found, can choose a random one. The ability to choose a random action has the advantage of emulating body-babbling, i.e. apparently random body movements that have the (hypothesized) purpose of learning the capabilities of the body in an environment [8]. Early in development, there are few experiences on which to draw, so random actions would be chosen more often, and later in development, it is more likely that an appropriate experience (and thus action) can be found. Additionally, with a small probability, the robot may still choose a random action as this may help move out of "local minima", and potentially discover new, more salient experiences.
To choose an action based on experience, the robot first examines the experience landscape for similar experiences near the current one. That is it finds a *candidate experience* with the shortest information distance to the current one. The next action that was executed following that experience is a *candidate action* to be executed next.
The candidate experience is chosen with a probability proportional to that experience's perceived *value* in terms of the stored value attributes (see section III-B above). The

[1] Vision sensors here are built by subdividing the visual field into regions and taking average colour values over each region at each timestep. In these experiments a 6x6 grid is used taking the average of the red channel only.

[2] Auditory channels were not used in the examples discussed.

[3] Note that sensor data is not being stored to build the interaction history, only the time-evolving probability distributions from which joint entropy can be estimated are stored.

[4] While this is probably not the most sophisticated model for acting, it is at least tractable.

exact nature of the calculation of *value* is dependent on the nature of the drives and motivations ascribed to the agent. For these experiments we use an internal variable that increases whenever a ball or human face is seen, but decays over time. This is explained in more detail in section *V-C*.

Finally, we introduce a feedback process that evaluates the result of any action taken in terms of whether there was an *increase in value* after the action was executed, and then adjusts the stored value attributes of the candidate experience, from which the action was derived, up or down accordingly. Closing of the perception-action loop in this way with feedback together with growth of the experiential metric space, results in the construction of modified behaviour patterns over time. This can be viewed as ontogenetic development, that is, as a process of change in structure and skills through embodied, structurally coupled interaction [9].

## IV. Geometry of Experience

In previous papers [7], [10]–[12] the authors have developed a mathematical geometry of experience that uses Shannon information theory [13] to place experience on a metric space as well as to compare sensorimotor experience using trajectories through projected sensor and motor spaces. The basis is the information metric [14], a measure of the "distance", in terms of information, between two random variables. We use the measure to compare sensorimotor experience over time and across modalities and the following is a brief overview of the relevant aspects.

### A. Information Distance

An agent situated and acting in an environment will have many external and internal sensory inputs any of which can be modeled as random variables changing over time. For any pair of sensors $\mathcal{X}$ and $\mathcal{Y}$ the *conditional entropy* $H(\mathcal{X}|\mathcal{Y})$ of $\mathcal{X}$ given $\mathcal{Y}$ is the amount of uncertainty that remains about the value $\mathcal{X}$ given that the value of $\mathcal{Y}$ is known.

$$H(\mathcal{X}|\mathcal{Y}) = -\sum_{x \in X} \sum_{y \in Y} p(x,y) \log_2 \frac{p(x,y)}{p(y)},$$

where $p(x,y)$ is given by the joint distribution of $\mathcal{X}$ and $\mathcal{Y}$.[5]

The *information distance*[6] between $\mathcal{X}$ and $\mathcal{Y}$ is then given by

$$d(\mathcal{X},\mathcal{Y}) = H(\mathcal{X}|\mathcal{Y}) + H(\mathcal{Y}|\mathcal{X}).$$

[5]We assume approximate local stationarity of the joint distribution of random variables representing the sensorimotor variables over a temporal window and that this can be estimated closely enough by sampling the sensorimotor variables.

[6]This satisfies the mathematical axioms for a *metric*:
1. $d(\mathcal{X},\mathcal{Y}) = 0$ if and only if $\mathcal{X}$ and $\mathcal{Y}$ are equivalent.
2. $d(\mathcal{X},\mathcal{Y}) = d(\mathcal{Y},\mathcal{X})$ (symmetry)
3. $d(\mathcal{X},\mathcal{Y}) + d(\mathcal{Y},\mathcal{Z}) \geq d(\mathcal{X},\mathcal{Z})$ (triangle inequality).

### B. Time-Horizon

Consider any sensor variable $\mathcal{X}$, beginning from a particular moment in time $t_0$ until a later moment $t_0 + h$ ($h > 0$), we regard the sequence of values $x(t_0), x(t_0+1), \ldots, x(t_0+h-1)$ taken by an information source $\mathcal{X}$ as time-series data from a new random variable $\mathcal{X}_{t_0,h}$, the *sensorimotor variable with temporal horizon $h$ starting at time $t_0$*

With this definition and that of information distance, we can then compare any sensorimotor variables over the same sized time-horizons, whether from the same sensor at different times, different sensors at the same time or, indeed, different sensors at different times.

### C. Experience Metric

We formalize an agent's *experience* from time $t$ over a temporal horizon $h$ as $E(t,h) = (\mathcal{X}^1_{t,h}, \ldots, \mathcal{X}^N_{t,h})$ where $\mathcal{X}^1, \ldots, \mathcal{X}^N$ is the set of all sensorimotor variables available to the agent. We can then define a metric on experiences of temporal horizon $h$ as

$$D(E,E') = \sum_{k=1}^{N} d(\mathcal{X}^k_{t,h}, \mathcal{X}^k_{t',h}),$$

where $E = E(t,h)$ and $E' = E(t',h)$ are two experiences of an agent and $d$ is the information distance (see [7], [10]).

## V. Experiments

We describe two experiments that explore the possibilities of the model of interaction history discussed. The first evaluates the veracity of the experience space by examining the ability of the model to predict future states of the world with reference only to the metric space of experience. The second shows early steps in using the model to play an interaction game, "peekaboo", with a human partner.

### A. Experiment 1: History-based Prediction

Given a robot[7] acting in an environment, how well can it predict future events based on its recent history of experience?

In this experiment the architecture was simplified, removing the developmental feedback loop, to examine the efficacy of using the metric space of experience to locate similar experiences. Two conditions were examined: in the first, the head stayed still while the ball was moved, and in the second a reactive process allowed the head to follow the ball.[8]

The *position* of the ball at the end of each experience was stored with the experience as a value attribute, and the *predicted future position* of the ball was given by the

[7]See Fig. 2. The robot used in this and all other experiments is the Sony AIBO ERS-7. Robot control programming was achieved using URBI-Universal Real-time Behaviour Interface [15].

[8]Simple colour based visual processing allowed the position of a pink ball in the visual field to be located as an $(X, Y)$ position, and the head would reactively move to centre that position.

Fig. 2. Sony Aibo ERS-7, Left: with pink ball, Right: hiding head while playing ”peekaboo”. The camera vision is partially obscured by the arm.

attributes stored with the experiences following the candidate (most similar previous) experience.

It is important to note that, the robot is not matching current ball position with previous ball position, rather we use all sensory and motor variables as information sources to detect similarity between experiences, and then use the tagged ball position to give the experimenter an indication as to how well the experience was chosen. For verification purposes a path is drawn on the display of the robot's visual field during operation, indicating the predicted future path.

### *B. Experiment 1: Results and Discussion*

In Fig. 3, we show a sequence of images from one trial from the first condition where the robot was passive while the ball was moved. The sequence lasts just over 4 seconds and consists of approximately 40 timesteps (1 timestep$\sim$100ms) and 8 experiences (experience granularity[9] of 5) [10].

In the sequence shown and others, the robot required very few examples of a sequence (usually one) before the appropriate experience could be located. This demonstrates that the information distance measure is capable of placing subjectively similar experiences (to an external observer) near to each other in the experience space. However, it was found that while the path of the ball could be predicted fairly well early on in the sequence, later on, as the choice of experiences grew, the candidate experience chosen was not always the most appropriate.

As an illustration of the problem, consider the eighth image in Fig. 3, here the predicted path from the candidate experience corresponds to the half circle that the ball has just been through (rather than the half-circle it is just about to go through, as in the other images). The candidate experience chosen is informationally close to another experience half

[9]Experience granularity denotes the number of timesteps between end-points of successive experiences. A granularity of 1 would store an experience of *horizon* timesteps at every timestep.

[10]Images are saved asynchronously at a rate of approx. 4 per second.. There were approximately 73 experiences at a granularity of 5 timesteps between experiences (about 38 seconds of activity) before the ones shown. Before the images shown, the ball was moved from left to right 4 times and in a clockwise circle once.

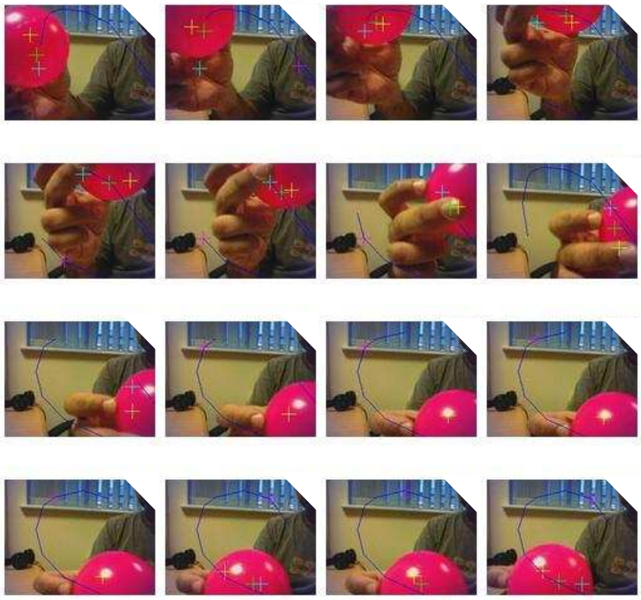

Fig. 3. Series of consecutive images from the Aibo camera showing ball path prediction using a sensorimotor interaction history. The robot does not move its head in this sequence. Images are sequential left to right and top to bottom and 147 images (73 experiences horizon length 40) precede these. The line shows the path prediction for 10 timesteps ahead. The crosses are from various methods for ball detection, only one of these was actually used as sensory input. Horizon=40, Number of Bins=5, Experience granularity=5 timesteps. Images captured approximately once every 2-3 timesteps.

a cycle back in time that may have been more appropriate, and the fact that the two possible experiences correspond to motions of the ball from opposite sides of a circle contributes to their being “recoding equivalents”[11], only differing in phase. Clearly, one solution to the issue is to provide the mechanism with more information, for instance from proprioception, with which to distinguish experience. The experiment is artificially hampered due there being no motor, active, component to the interaction.

Fig. 4 shows a series of images showing the path prediction in the second condition, where the robot was actively following at the ball with its head. The ball path is generally a small loop starting at and finishing near the centre of the image. This is to be expected as, since we are plotting just the position of the ball *within the image* then, this cannot describe the absolute position of the ball in space. To better assess the result of the experiment, we would need to have the predicted position of the head rather than the ball. Further work, will look into the predictive capabilities of the method with regard to the robot acting as a whole.

### *C. Experiment 2: Sensorimotor contingencies in an interaction game - Peekaboo*

The purpose of this experiment was to investigate whether the development of an enactive interaction history in a

[11]That is, are a small information distance apart.

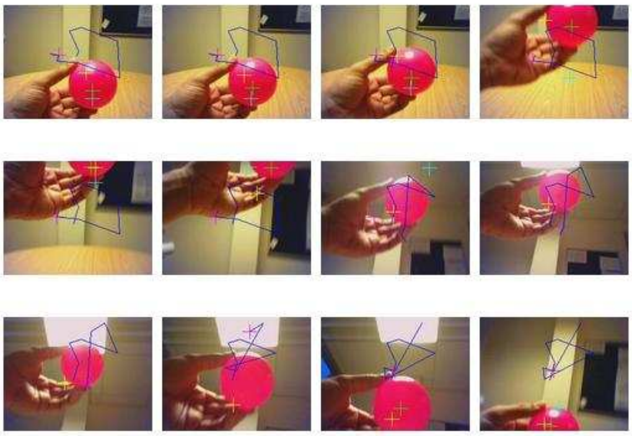

Fig. 4. Series of consecutive images (left to right, top to bottom) from the Aibo camera showing ball path prediction using a sensorimotor interaction history. The robot's head reactively follows the ball in this sequence. Images are sequential left to right and top to bottom and the sequence is approximately 3 seconds long. (See text and Fig. 3 for further notes). Horizon=20, Number of Bins=2, Experience granularity=4 timesteps.

robot could be used for the robot to act appropriately in an interaction that required following a spatio-temporally structured set of "rules", that when followed result in high motivational value. The full architecture was used, with the action feedback loop modifying potential future interaction.

The simple interaction game of *peekaboo* played between adults and babies or young children was taken as a model. The game consists of a repeated cycle of an initial contact, disappearance, reappearance, and acknowledgment of renewed contact [16]. Bruner and Sherwood suggest that the peekaboo game may provide scaffolding for further interaction and learning [16] and as such is useful in studying the development of interaction capabilities in a robot in a social environment.

Bruner and Sherwood also suggest that the peekaboo game itself may emerge from the exploitation of innate tendencies or motivations in the child and we model important aspects of potential precursors to this game as actions, drives and motivations of the cognitive model of the robot. Specifically, the robot gains "pleasure" (increase in internal variable 1) in seeing a face, however if the face is lost, it has a rising "expectation" (internal variable 2) of seeing the face again, and the "pleasure" in seeing the face at a later time is increased by the value of that expectation. The atomic actions implemented for the selection mechanism were: 1) move head up, 2) down, 3) left, 4) right and 5) hide/reveal head.
The robot is preprogrammed with abilities to recognize a generalized face[12] and this yields sensory variables indicating the position of the face in the visual field.

[12] Implemented using Intel OpenCV HAAR Cascades [17].

### D. Experiment 2: Results and Discussion

Thus far we have completed a basic feasibility study with one of the authors interacting with the robot. The results tend to show that the robot, after a period of random movement does start to engage in repeated cycles of behaviour, Fig. 5. If the robot were not to hide its face, it would have long periods of seeing the face which do not result in high motivational value (internal variable 1), instead the robot generates intermittency in seeing the face by hiding its own face resulting in high motivational value when the face is next seen. This often includes cycles of hiding and revealing the face, as shown in Fig. 5.

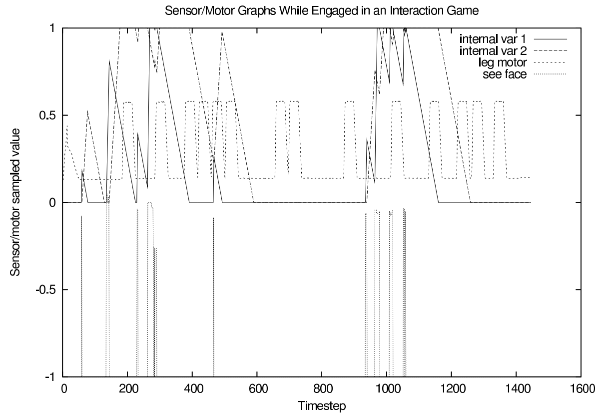


Fig. 5. Time series of motor and sensor values showing engagement of robot in peekaboo game. The bottom part of the graph shows when the face is seen and the two internal variables are shown varying in response to this. The peaks in the leg motor trace indicate when the robot is hiding its head with its foreleg.

Fig. 6 shows the value assigned to experiences and how these change over time. It is clear that only a few experiences are regularly selected and thus modified over time, increasing and decreasing in value. The final metric space of experiences is depicted in Fig. 7, and indicates that the experience space has a consistent (non-random) structure with definite peaks that correspond to those few experiences that become present candidate experiences.

## VI. Discussion and Future Directions

The positive results from the experiments using the experience space to predict future experiences indicate that the method of information distance has the potential of forming the basis of an interaction history, particularly if the whole embodied experience of the robot is taken into account. However, mechanisms may be needed to disambiguate the experience in the space when there are many experiences to select from. Steps towards this are made in using "value" to test candidate experiences against each other, however, other mechanisms might be considered, e.g. finding exemplary experiences by grouping near experiences. Further, it is also

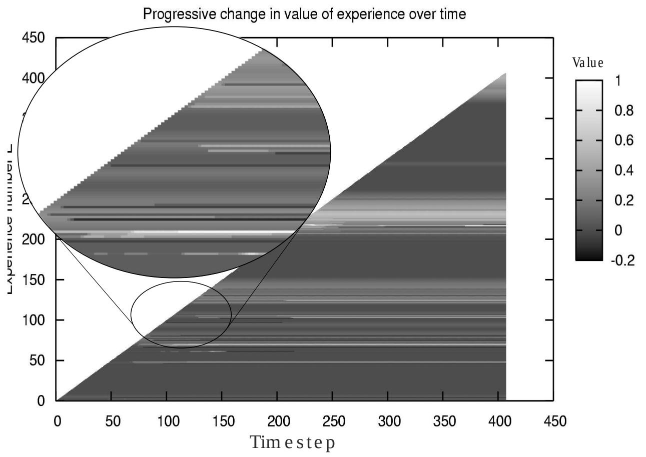


Fig. 6. Graph showing the "value" (as shade) assigned to experiences (on vertical axis), and how this progresses over time as values are changed while the robot actively reconstructs its experience space. The zoomed in region shows individual experiences changing in value. Note that the triangular shape is due to new experiences being added over time, and that most experiences do not change in their values.

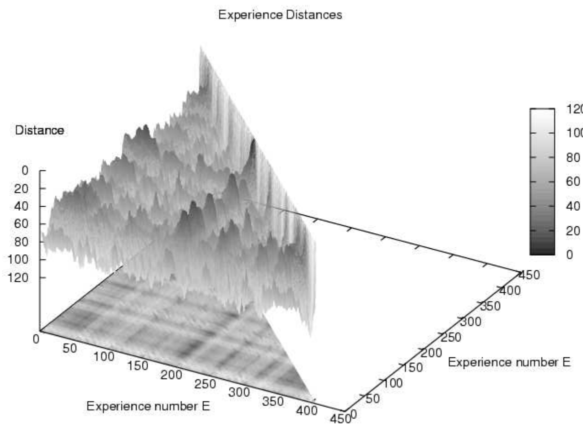


Fig. 7. Depiction of experience space at the end of the run shown in Fig. 6. The axes in the plane are the experiences being compared while the height indicated their experiential information distance. The black peaks are low information distances and indicate "similar" experiences. It is these experiences that provide candidates from which to select action.

clear that experiences of different time horizon sizes will be needed to anticipate experience on different timescales.

There are also good indications that the method of choosing action based on "value" can be useful in choosing between many potentially similar experiences, however, how this value is assigned and modified will need to be made more sophisticated in order to better assign credit to appropriate action, and to handle multiple, potentially conflicting, goals. Similar comments apply to the simple method of action selection. In a more complex environment an action selection mechanism that can deal with appropriate action in a particular context, and that can deal with parallel and temporally extended actions and behaviours would be needed.

Future directions for the research will include investigating more generic approaches to ascribing motivation to artificial agents in order to select experience and action, for example sensorimotor contingencies, drives for comfort (predictability of environment) or novelty (see for example [18]). Further work will be conducted on the peekaboo game as a testbed in which to study the development of interactive behaviour.

## References


[1] H. R. Maturana and F. J. Varela, *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston: New Science Library (Shambhala), 1987.

[2] W. J. Clancey, *Situated Cognition: On human knowledge and computer representations*. Learning in doing: Cognitive and compuational perspectives, Cambridge University Press, 1997.

[3] K. Dautenhahn and T. Christaller, "Remembering, rehearsal and empathy - towards a social and embodied cognitive psychology for artifacts," in *Two Sciences of the Mind: Readings in cognitive science and consciousness* (S. O'Nuallain and P. M. Kevitt, eds.), pp. 257–282, John Benjamins North America Inc., 1996.

[4] E. Tulving, *Elements of episodic memory*. Oxford: Clarendon Press, 1983.

[5] C. L. Nehaniv, K. Dautenhahn, and M. J. Loomes, "Constructive biology and approaches to temporal grounding in post-reactive robotics," in *Sensor Fusion and Decentralized Control in Robotics Systems II (September 19-20, 1999, Boston, Massachusetts), Proceedings of SPIE* (G. T. McKee and P. Schenker, eds.), vol. 3839, pp. 156–167, 1999.

[6] E. T. Rolls, *The Brain and Emotion*. Oxford University Press, 1999.

[7] C. L. Nehaniv, "Sensorimotor experience and its metrics," in *Proc. of 2005 IEEE Congress on Evolutionary Computation*, 2005.

[8] A. Meltzoff and M. Moore, "Explaining facial imitation: a theoretical model," *Early Development and Parenting*, vol. 6, pp. 179–192, 1997.

[9] M. Lungarella, G. Metta, R. Pfeifer, and G. Sandini, "Developmental robotics: A survey," *Connection Science*, vol. 15, no. 4, pp. 151–190, 2004.

[10] C. L. Nehaniv, N. A. Mirza, K. Dautenhahn, and R. te Boekhorst, "Extending the temporal horizon of autonomous robots," in *Proc. of the 3rd International Symposium on Autonomous Minirobots for Research and Edutainment (AMiRE2005)* (K. Murase, K. Sekiyama, N. Kubota, T. Naniwa, and J. Sitte, eds.), pp. 389–395, Springer, 2006.

[11] N. A. Mirza, C. Nehaniv, R. te Boekhorst, and K. Dautenhahn, "Robot self-characterisation of experience using trajectories in sensory-motor phase space," in *Proc. of Fifth International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems (EpiRob2005)*, pp. 143–144, Lund University Cognitive Studies, 2005.

[12] N. A. Mirza, C. L. Nehaniv, K. Dautenhahn, and R. te Boekhorst, "Using sensory-motor phase-plots to characterise robot-environment interactions," in *Proc. of 6th IEEE International Symposium on Computational Intelligence in Robotics and Automation (CIRA2005)*, pp. 581–586, 2005.

[13] C. E. Shannon, "A mathematical theory of communication," *Bell Systems Technical Journal*, vol. 27, pp. 379–423 and 623–656, 1948.

[14] J. Crutchfield, "Information and its metric," in *Nonlinear Structures in Physical Systems - Pattern Formation, Chaos and Waves* (L. Lam and H. Morris, eds.), pp. 119–130, New York: Springer-Verlag, 1990.

[15] J.-C. Baillie, "Urbi: Towards a universal robotic low-level programming language," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems 2005*, 2005. http://www.urbiforge.com.

[16] J. S. Bruner and V. Sherwood, "Peekaboo and the learning of rule structures," in *Play: Its Role in Development and Evolution* (J. Bruner, A. Jolly, and K. Sylva, eds.), pp. 277–285, New York: Penguin, 1975.

[17] OpenCV, "Open computer vision library." http://sourceforge.net/projects/opencvlibrary/ (GPL), 2000.

[18] P.-Y. Oudeyer, F. Kaplan, V. V. Hafner, and A. Whyte, "The playground experiment: Task-independent development of a curious robot," in *Proc. of the AAAI Spring Symposium on Developmental Robotics, 2005* (D. Bank and L. Meeden, eds.), pp. 42–47, 2005.